Beginning of 02/28/2025

More on Markov Chains

A process $\{X_t\}$ is Markovian if $\mathbb{P}(X_{t+1} \mid X_t, \cdots, X_0) = \mathbb{P}(X_{t+1} \mid X_t)$, i.e., given the present state, the rest of the past is irrelevant for predicting the future. Notation-wise, we define the following:

(1) $P = \{P_{i,j}\}$ is the state transition matrix where $P_{i,j} = \mathbb{P}(X_t = j \mid X_{t-1} = i)$.

(2) x_0 is the initial distribution, and x_t is the distribution at time t.

(3) Assuming P is erdogic, the steady state distribution π exists and is unique. Then $\pi P = \pi$ and $\lim_{t\to\infty} x_0 P^t = \pi$.

Last time we discussed a random walk on graph: given G, let $P_{u,v} = 1/2d(u)$ [degree] and $P_{u,u} = 1/2$. Then the steady state distribution satisfies $\pi_u = d(u)/2|E|$. Furthermore, for any edge (u, v),

$$\pi_u P_{u,v} = \pi_v P_{v,u} = \frac{1}{4|E|},$$

which implies that in the steady state, the chain is transitioning via any directed edge is uniform as u, v can be arbitrary. Specifically, in an idealized scenario where we *assume stable distribution*, if we fix an edge (u, v), the expected number of steps that we revisit (u, v) for a second time *since we first visited* (u, v) is $4|E|$.
What if we deal with the distributions x_t instead of π? Wald's **renewal theorem** shows this intuitive result holds. More formally, let Q be the time for walk starting at v to walk along $u \to v$. This indicates one **renewal** (i.e. revisit). Note this is well-defined since Markov chains are memory-less. Let $N(t)$ be the number of renewals (i.e. revisits) in time t. Then,

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{4|E|} = \frac{1}{\mathbb{E}[Q]} \qquad \text{and} \qquad \mathbb{E}[\text{time to revisit } (u \to v) \mid \text{state } v] = 4|E|.$$

More definitions.

Definition: Hitting Time

Define $h_{u,v} = \mathbb{E}[\text{time to reach } v \mid \text{state } u]$ to be the **hitting time** from u to v, and $h_{u,v} + h_{v,u}$ the **commute time** between u and v.

Immediately, since the expected time of $v \rightsquigarrow (u \to v)$ is $4|E|$, the expectation of $u \rightsquigarrow v \rightsquigarrow u$ is certainly no more, so

$$h_{u,v} + h_{v,u} \leqslant 4|E| \qquad \text{for all } (u, v) \in E.$$

Definition: Cover Time

... is defined to be $\mathbb{E}[\text{time to visit all of } V \mid \text{ start at } u]$.

We claim that regardless of starting state, the cover time is bounded by $4m(n-1)$. Given a graph G, we restrict our attention to its MST and convert it to a bidirectional graph, adding both $(u \to v)$ and $(v \to u)$ to the edge set E (with abuse of notation). The modified MST is then Eulerian, so we can find a Eulerian tour via for example DFS.

$$\text{cover time} \leqslant \sum_{(u\to v)\in\text{tour}} h_{u,v} = \sum_{(u,v)\in\text{tree}} [h_{u,v} + h_{v,u}] \leqslant \sum_{(u,v)\in\text{tree}} 4m = 4m(n-1).$$

Logspace Algorithms

Consider the $s-t$ connectivity problem on an undirected graph. Assume the input is given for free. Usually, to solve the problem we need *extra* memory depending on the graph size as we need to maintain data structures such as stacks or queues. However, below we propose an algorithm which takes much less *extra* memory:

- Start at s and run a random walk for $4n^3$ steps;
- If t is reached, output YES; else output NO.

Essentially all it needs is a random number generator and access to the already provided graph. This algorithm belongs to **randomized logspace** (RL). If the answer is YES it is clearly correct, so we just need to bound the false negative probability. Using the identity cover time $\leqslant 4m(n-1) \leqslant 2n^3$, we see

$$\mathbb{E}[\text{time to reach } t \mid \text{start} = s] \leqslant \text{cover time} \leqslant 2n^3.$$

Therefore $\mathbb{P}[t \text{ not reached in } 4n^3 \text{ steps}] \leqslant 1/2$, bounding false negative probability by $1/2$. (And of course we can make this bound smaller if needed.)

It was shown in 2004 that L (logspace) = RL, i.e., one does not need randomized algorithm to achieve this effect.

How Fast Do Distributions Converge?

We would like to know how fast x_t converges sufficiently close to the stable distribution π. But first, some definitions. Given two distributions p, q, the **total variation distance** is

$$\mathrm{TVD}(p,q) = \frac{1}{2}\sum_s |p_s - q_s|.$$

Given a Markov chain, let

$$p_x^t(i) = \mathbb{P}[\text{state} = i \text{ at time } t \mid \text{state} = x \text{ at } t = 0]$$

and

$$\Delta_x^t = \mathrm{TVD}(p_x^t, \pi)$$

so that Δ_x^t measures how close the distribution at time t if starting at x is to our goal π. Finally, Δ_t is the worse possible deviation at time t, $\max_x \Delta_x^t$. Finally,

Definition: Mixing Time

Given $\epsilon > 0$, the **mixing time** $\tau(\epsilon)$ is the first time when the TVD is ϵ-approximate of π:

$$\tau(\epsilon) = \min_t\{\Delta(t) \leqslant \epsilon\}.$$